



Salient contour extraction from complex natural scene in night vision image



Jing Han, Jiang Yue, Yi Zhang, Lian-fa Bai *

Jiangsu Key Laboratory of Spectral Imaging and Intelligent Sense, Nanjing University of Science and Technology, Nanjing 210094, China

HIGHLIGHTS

- We analyze LLL and infrared images with natural scene base on nCRF mechanisms.
- An optimized compound modulation method extracts contour from night vision image.
- MFC weighting inhibition model improves surrounding inhibition.
- FC facilitation model achieves the connection of discontinuous contours.
- A multi-scale iterative attention method accomplishes dynamic modulation process.

ARTICLE INFO

Article history:

Received 14 June 2013

Available online 4 January 2014

Keywords:

LLL image

Infrared image

Complex natural scene

Contour extraction

Compound modulation

ABSTRACT

The theory of center-surround interaction in non-classical receptive field can be applied in night vision information processing. In this work, an optimized compound receptive field modulation method is proposed to extract salient contour from complex natural scene in low-light-level (LLL) and infrared images. The kernel idea is that multi-feature analysis can recognize the inhomogeneity in modulatory coverage more accurately and that center and surround with the grouping structure satisfying Gestalt rule deserves high connection-probability. Computationally, a multi-feature contrast weighted inhibition model is presented to suppress background and lower mutual inhibition among contour elements; a fuzzy connection facilitation model is proposed to achieve the enhancement of contour response, the connection of discontinuous contour and the further elimination of randomly distributed noise and texture; a multi-scale iterative attention method is designed to accomplish dynamic modulation process and extract contours of targets in multi-size. This work provides a series of biologically motivated computational visual models with high-performance for contour detection from cluttered scene in night vision images.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Contour extraction plays an important role in night vision images (low-light-level (LLL) and infrared images) understanding. Most applications of night vision target detection and recognition are in outdoor scene, which contains a great deal of natural texture like grass or foliage. The traditional edge detection operators, such as Canny, retain many non-contour boundaries. So how to eliminate uninteresting local boundaries engendered from textures and maintain the integrated contour are the major problems in night vision contour extraction.

Many solutions to extract contour from complex scene are proposed. In those methods, contour extraction through non-classical receptive field model based on the mechanism of biological vision [1] is distinguished by remarkable performance in the high-quality

visible images. The area beyond classical receptive field (CRF) of neurons in primary visual cortex (V1), namely non-classical receptive field (nCRF), can modulate the stimuli response in CRF [2–10]. It has been found that the majority of the surround modulation is suppressive, which makes the isolated contour more significant than homogeneous background. Grigorescu et al. [2,3] proposed a series of related nCRF inhibition models to extract contour, including isotropic and anisotropic inhibition models [4], which utilize surround suppression in center direction to weaken natural texture. Giuseppe et al. [5] developed a multi-resolution contour detection method based on Bayesian denoising and a surround inhibition technique. The model built by Sang et al. [6] with butterfly-shaped subregion, weighting inhibition by the difference of stimulus-direction between CRF and nCRF, can decrease the impact of collinear suppression. Ursino et al. [7,8] introduced level attention mechanism to gain contour in different scales. Zeng et al. [9,10] improved the butterfly-shaped inhibition to present a two-scale based contour extraction model and an orientation-selective inhibition model.

* Corresponding author. Tel.: +86 02584315247.

E-mail addresses: mrblf@163.com, njusthanjing@163.com (L.-f. Bai).

These approaches work well on suppressing background in visible images with high resolution, as is shown in Fig. 1, but they do not always perform well in LLL and infrared images. Compared with high-quality visible image, heavier noise and lower contrast in night vision images lead to inaccurate and cracked contour, and it is difficult to suppress natural texture. So in this paper, we design a reasonable inhibition model for contour extraction in night vision image, namely multi-feature contrast (MFC) weighted inhibition model.

On the other hand, inhibition weakens the strength of contour response. Contours in night vision image are prone to disconnect due to imaging characteristics, surround inhibition and noise interference, which will affect the subsequent target recognition.

Some physiological findings [11] have shown that modulations from nCRF are both inhibition and facilitation. Tang et al. [12,13] added collinear excitation in nCRF inhibition model, which calculates facilitation based on curvature in grouping properties to enhance the response of contour points. In the aspect of contour grouping, Sha et al. [14] proposed a significance calculation method based on curvature change and developed a grouping algorithm to extract remarkable curve. Geisler et al. [15] testified that contour grouping in natural scene corresponds with the co-circularity rule by studying the statistical properties of contours in natural image, and carried out a maximum likelihood local grouping function for contour. Elder et al. [16] studied the statistical properties of contour perceptual organization in Gestalt rules, including proximity, continuity and similarity, and suggested a Bayesian ratio model to detect contour. Wang et al. [17,18] provided a ratio-contour method to extract significant closed contour based on graph model. Guy et al. [19–21] developed a global cognitive model, namely tensor voting, to infer significant structures in image by tensor analysis and neighbor voting on each point. These contour grouping models devote to organize local segments to global contour information.

However, it is difficult to realize accurate contour connection through single grouping rule. In this paper, more contour grouping constraints are introduced to consider facilitation mechanism. We propose a fuzzy connection (FC) facilitation model to enhance weak contours and connect discontinuous contours to keep the integrated salient contours.

A multi-scale iterative attention method proposed in this paper, combining the MFC weighted inhibition model and FC facilitation model, realizes the optimized compound modulation. It primarily solves the problems of surrounding suppression and contour connection in LLL and infrared images.

The rest of this paper is organized as follows. Section 2 details our method of salient contour extraction in night vision images. Section 3 discusses the analysis of presented models. Section 4 gives the practical application and performance of our method in LLL and infrared images in nature scenes. Finally, Section 5 summarizes the main conclusions of this paper and the future works.

2. Visual model of salient contour extraction in night vision

2.1. nCRF model

Two-dimensional Gabor function is used to describe the receptive field profile of simple cells in visual cortex. The response modulus of odd-even pair of Gabor filters, called Gabor energy, can effectively models fundamental characteristics of typical complex cells, which can be seen as local orientation energy operators [22]. The maximum response of complex cell is used to accurate positioning edge and line in graphics. So Gabor energy is chosen to simulate the activities of complex cell. Two-dimensional Gabor filter can be expressed as follow.

$$g_{\lambda\sigma\varphi}(x, y; \theta) = \exp\left(-\frac{\tilde{x}^2 + \gamma^2\tilde{y}^2}{2\sigma^2}\right) \cdot \cos\left(2\pi\frac{\tilde{x}}{\lambda} + \varphi\right) \quad (1)$$

where $\tilde{x} = x \cos \theta + y \sin \theta$, $\tilde{y} = -x \sin \theta + y \cos \theta$, in which θ is the preferred orientation of CRF; φ is phase offset; $g_{\lambda\sigma\varphi}(x, y; \theta)$ with $\varphi = 0$ and $\varphi = \pi/2$ represent the even and odd Gabor filter respectively; γ is the spatial aspect ratio, which determines the eccentricity of the Gaussian envelope; λ is wavelength; σ is the standard deviation of Gaussian factor, which decides the area of the Gabor simulated CRF; σ/λ represents the spatial frequency bandwidth.

Response of simple cells $r_{\lambda\sigma\theta\varphi}$ is given as the convolution of Gabor function and image I . Neuronal response to stimulus in CRF $E_{\sigma}(x, y; \theta_i)$ is defined as the combination of orthogonal pair of simple cells' responses.

$$r_{\lambda\sigma\varphi}(x, y; \theta) = I(x, y) * g_{\lambda\sigma\varphi}(x, y; \theta) \quad (2)$$

$$E_{\sigma}^2(x, y; \theta_i) = r_{\lambda\sigma 0}^2(x, y; \theta_i) + r_{\lambda\sigma \pi/2}^2(x, y; \theta_i) \quad (3)$$

where Gabor energy of CRF is divided into N_{θ} directions: $\theta_i = i\pi/N_{\theta}$, $i = 0, 1, \dots, N_{\theta} - 1$. According to [9] we set $N_{\theta} = 12$, $\gamma = 0.5$, $\sigma/\lambda = 0.56$. σ is a variable in our experiments. Output response of CRF is the maximum of $E_{\sigma}(x, y; \theta_i)$ over all orientations.



Fig. 1. Results of contour extraction in natural images. The columns from left to right represent the original images, outputs of canny, butterfly-shaped inhibition [6], orientation-selective inhibition [10] and two-scale inhibition [9] models, respectively.

$$E_\sigma(x, y) = \max\{E_\sigma(x, y; \theta_i) | i = 0, 1, \dots, N_\theta - 1\} \quad (4)$$

Physiological studies show that visual cortex cells are inhibitory principally. The surround inhibition is used to segment inhomogeneous region and highlight boundary. In those previous vision-based models [5,6,10], a special distance weighting function $\omega_\sigma(x, y)$ is usually used to describe the spatial structure of surround inhibition. The weighting function $\omega_\sigma(x, y)$ is defined as a half-wave rectified and L1-normalized difference of two concentric Gaussian functions $DOG_{\sigma, k\sigma}(x, y)$.

$$DOG_{\sigma, k\sigma}(x, y) = \frac{1}{\sqrt{2\pi(k\sigma)^2}} \exp\left(-\frac{x^2 + y^2}{2(k\sigma)^2}\right) - \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (5)$$

$$\omega_\sigma(x, y) = |DOG_{\sigma, k\sigma}(x, y)|^+ / \| |DOG_{\sigma, k\sigma}(x, y)|^+ \|_1, \quad |z|^+ = \begin{cases} 0 & z < 0 \\ z & z \geq 0 \end{cases} \quad (6)$$

The inhibition term $t_\sigma(x, y; \theta_i)$ is defined as the distance weighted local orientation energy $E_\sigma(x, y; \theta_i)$ in annular surround of each pixel.

$$t_\sigma(x, y; \theta_i) = \omega_\sigma(x, y) * E_\sigma(x, y; \theta_i) \quad (7)$$

Here σ and $k\sigma$ determine the area of CRF and nCRF respectively. We set $k = 3$ in our experiments, because that nCRF diameter may reach 2–5 times than CRF diameter. $|DOG|^+$ ensures that $\omega_\sigma(x, y)$ only works on nCRF region as is shown in Fig. 2. Radius r_c of CRF area S_c is got by setting $DOG_{\sigma, k\sigma}(x, y) = 0$. Thus $r_c = \sqrt{2 \ln k / (k^2 - 1)} \cdot k\sigma$ and the radius of nCRF area S_{nc} is kr_c .

2.2. New method for salient contour extraction in night vision images

The new method in this paper is shown in Fig. 3. The presented multi-feature contrast weighted inhibition model and fuzzy connection facilitation model is executed synchronously to suppress background texture and connect of discontinuous contour. Input the results of inhibition and facilitation into the multi-scale iterative attention method, we accomplish dynamic modulation and extract contours of targets in multi-size. It performs well in salient contour extraction from night vision images in clutter nature scenes.

2.3. MFC inhibition model

The isotropic annular suppression structure mentioned above may cause mutual inhibition among inhomogeneous regions, which will result in weak contour. Butterfly-shaped and orientation-selective models [6,9,10] weight inhibition by orientation difference of CRF and nCRF to reduce the response in homogeneous region and maintain the response in inhomogeneous region. But they are susceptible to noise and easy to get inaccurate orientation in center and surround. So it is difficult for single feature to

evaluate the homogeneity between center and surround, and is hard to generate precise suppression in night vision images.

Experiments have shown that the inhibitory degree depends on the difference between optimal response orientation (preferred orientation) and stimulus orientation in neurons. Modulatory effect is strongest when CRF and nCRF have similar preferred orientation [23–25] and it weakens with the increasing of the orientation difference [23,24,26,27]. Gray spatial frequency [25] and contrast [28] play similar roles. Thus the inhibition is strongest when accordant features in CRF and nCRF, it decreases or disappears when there are differences in center and surround. So we introduce multidimensional features analysis to present a multi-feature contrast (MFC) weighted inhibition model.

Compared with high-quality image, signal noise ratio is low in LLL image and boundaries are blurry in infrared image. Contour influenced by serious noise in LLL image gives no obvious orientation. But local spatial frequency around contour is more prominent than surroundings in global noise. Blurry contour in infrared image also has weak orientation response. But local contrast of contour is higher than surroundings. So we design a features selection unit, including spatial frequency SF , contrast CON , co-occurrence statistics of homogeneity $cooH$ and correlation $cooCOR$ (Appendix A is the computing of these features), to detect comprehensive differences between nCRF and CRF. The feature vector of CRF is defined as $\mathbf{x} = [SF_c, CON_c, cooH_c, cooCOR_c]^T$ and feature vector of nCRF is defined as $\mathbf{y} = [SF_{nc}, CON_{nc}, cooH_{nc}, cooCOR_{nc}]^T$. Here, SF_c , SF_{nc} , CON_c , CON_{nc} , $cooH_c$, $cooH_{nc}$, $cooCOR_c$ and $cooCOR_{nc}$ are normalized spatial frequency, contrast, co-occurrence statistics of homogeneity and correlation in regions of CRF S_c and nCRF S_{nc} respectively. MFC value changes with the angle θ_Δ and Euclidean distance d_Δ , which represent the direction and distance between feature vectors in feature space. θ_Δ and d_Δ measure the contrast of multi-feature in center and surround. The MFC weighting is defined as follow.

$$\omega_{mfc}(x, y) = \exp\left(-\frac{(1 - \cos \theta_\Delta)^2 + \varepsilon_1 d_\Delta^2}{2\sigma_\Delta^2}\right) \quad (8)$$

$$\cos \theta_\Delta = \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\| \cdot \|\mathbf{y}\|}, \quad d_\Delta = \|\mathbf{x} - \mathbf{y}\| \quad (9)$$

Here ε_1 controls the proportion of d_Δ and θ_Δ in multi-feature contrast. The standard deviation of Gaussian distribution σ_Δ determines the attenuation degree for MFC weighting $\omega_{mfc}(x, y)$ with multi-feature contrast.

In Fig. 4(c1, c2), we compare the differences of center and surround windows in small and large scales from the aspects of orientation, spatial frequency, contrast, co-occurrence statistics of correlation and homogeneity. The sizes of center and surround windows are 3×3 and 7×7 in small scale, 7×7 and 21×21 in large scale. We can see that it is difficult to distinguish the center and surround in weak contour regions by only orientation, the difference of orientation is small especially in large scale, but the other four features can discriminate better either in small scale or large scale.

The multi-feature can distinguish center and surround well. So the MFC weighting defined in Eq. (8) constructed by the distance and angle of multi-feature vectors can measure the difference between CRF and nCRF region effectively. In Fig. 4, it is evident that MFC value of contour point is outstanding (d_1 , d_2) in situations of strong noise (b1) and low contrast (b2). Through analyzing the multi-feature in center and surround, pixels with similar feature vectors in CRF and nCRF have small θ_Δ and d_Δ to gain high ω_{mfc} in homogeneous region. Whereas, contour pixels with different feature vectors get larger θ_Δ and d_Δ to generate relatively lower ω_{mfc} , which allows the MFC weighting of contour to stand out from background.

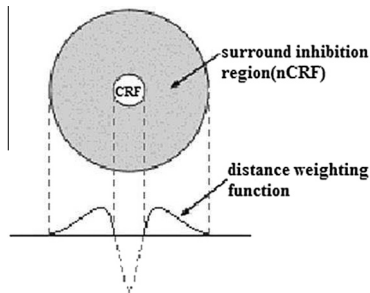


Fig. 2. The distance weighting function of surround inhibition.

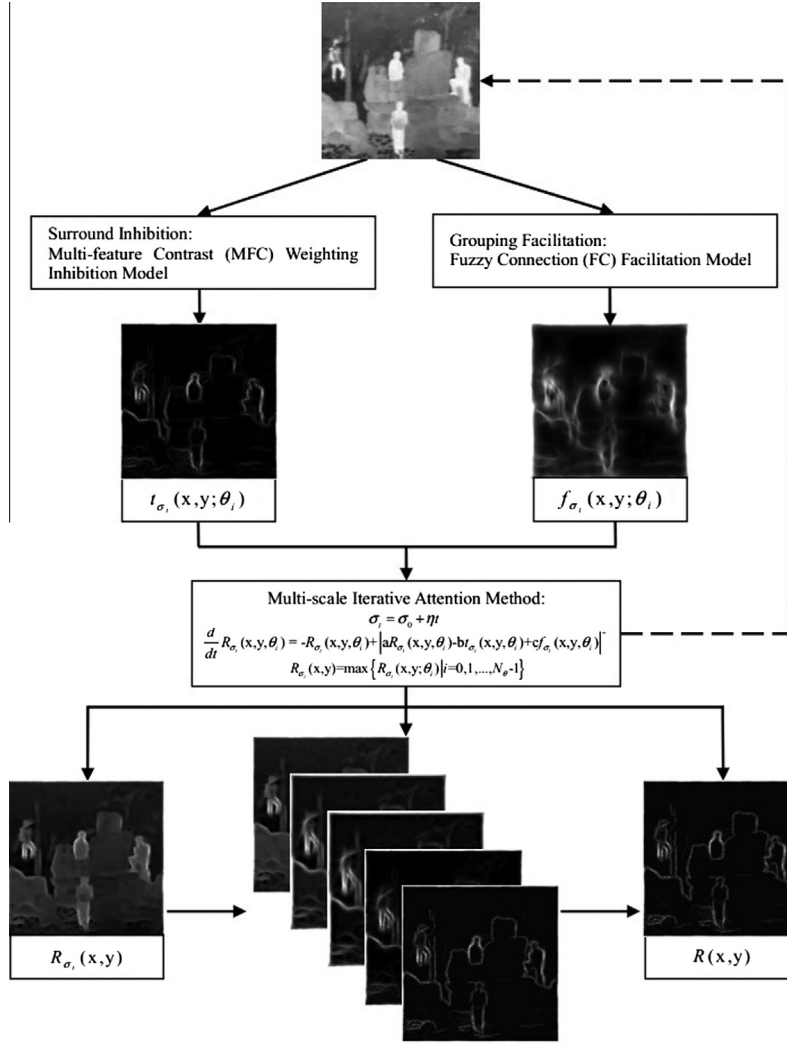


Fig. 3. General flowchart of the proposed method.

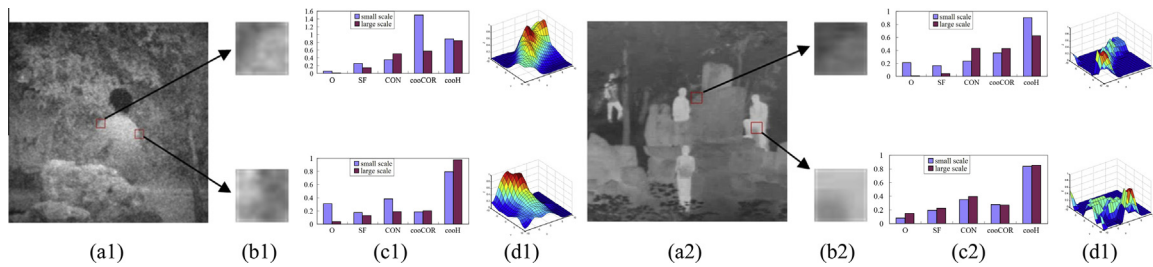


Fig. 4. Multi-feature contrast analysis. (a1) and (a2) original image, (b1) and (b2) enlarged view of red box area with 21×21 pixels, (c1) and (c2) center and surround differences by different features in small and large scales, (d1) and (d2) grid graph of value $1 - \omega_{mfc}$ in red box region. Parameters are set as $\epsilon_1 = 0.4$, $\sigma_{\Delta} = 0.5$, $\sigma = 3$ for LLL image and $\sigma = 2.5$ for infrared image.

We also consider that inhibitory strength decreases with the increasing distance from nCRF pixels to CRF center $\omega_{\sigma}(x, y)$ and the orientation-selective mechanism $\sum_{j=1}^{N_{\theta}} \cos(\theta_i - \theta_j)$, the output of MFC inhibition $t_{\sigma}(x, y; \theta_i)$ is finally improved as follow.

$$t_{\sigma}(x, y; \theta_i) = \omega_{mfc}(x, y) \cdot \sum_{j=1}^{N_{\theta}} \cos(\theta_i - \theta_j) \cdot [\omega_{\sigma}(x, y) * E_{\sigma}(x, y; \theta_j)] \quad (10)$$

Here, θ_i and θ_j are the orientations of pixel in CRF and nCRF.

Compared with original inhibition $\omega_{\sigma}(x, y) * E_{\sigma}(x, y; \theta_i)$ in Eq. (7), the Eq. (10) $\omega_{mfc}(x, y) \cdot \sum_{j=1}^{N_{\theta}} \cos(\theta_i - \theta_j) \cdot [\omega_{\sigma}(x, y) * E_{\sigma}(x, y; \theta_j)]$ adds the contrast of orientation $\sum_{j=1}^{N_{\theta}} \cos(\theta_i - \theta_j)$ and MFC weighting $\omega_{mfc}(x, y)$ to gain the ability of improving surround inhibition. In homogeneous area the MFC weighting is large, the inhibition is strong. In inhomogeneous area the MFC weighting is much smaller, the inhibition is weak. So we can suppress most of the texture regions and keep the main contours.

Compared with orientation-selective mechanism in previous methods, MFC inhibition model has a higher robustness to apply

in night vision images by considering features of preferred orientation, spatial frequency, contrast and co-occurrence statistics. As is shown in Fig. 5, the MFC inhibition model can suppress noise and texture more effectively to segregate contour from background with less false boundaries. Furthermore, introducing multi-feature analysis can reduce mutual inhibition among contour elements to strengthen contour response.

2.4. FC facilitation model

In Fig. 5 there are still false boundaries and cracked contours in LLL and infrared images due to imaging quality and excessive surrounding suppression. How to balance the suppression of background and retention of contour is an essential problem. It is remarkable that contour grouping is dedicated to extract coherent structures in a field of randomly oriented elements. Thus, the principle of grouping can be used to strengthen contour response and organize local boundaries into integrated contour.

In Gestalt rule [7], visual system tends to group consistent image features to higher-level structures through some rules, including that curvature constancy is prone to maintain the original curvature in curve extending; low curvature tends to connect segments by curve with low curvature; local proximity is in favor of connecting segments with short distance.

Geisler et al. [12,15] provided a maximum likelihood local grouping function of co-circularity rule for contour grouping in natural scenes. The co-circularity rule requires that connective elements should be tangent to a circle. Thus optimal orientation of pixel A $\theta_i (0 \leq \theta_i < \pi)$ and direction of pixel B $\theta_j (0 \leq \theta_j < \pi)$ should follow Eq. (11) in Fig. 6. Here δ is the deviation, which can be set as π/N_θ .

$$\begin{cases} \theta_j = 2\theta_{AB} - \theta_i \pm \delta & 0 \leq 2\theta_{AB} - \theta_i < \pi \\ \theta_j = 2\theta_{AB} - \theta_i + \pi \pm \delta & 2\theta_{AB} - \theta_i < 0 \\ \theta_j = 2\theta_{AB} - \theta_i - \pi \pm \delta & 2\theta_{AB} - \theta_i \geq \pi \end{cases} \quad (11)$$

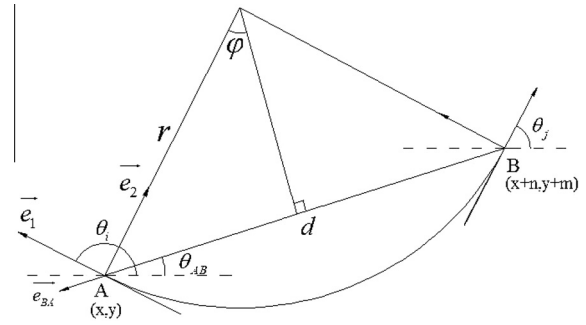


Fig. 6. The geometrical relationship of co-circularity constraint.

However, co-circularity constraint is not comprehensive according to Gestalt rule. We propose a fuzzy connection (FC) facilitation model for contour integration. An optimized grouping constraint, combining the co-circularity rule with the preference of curvature constancy, low curvature and local proximity in visual perception, is designed to connect local segments in consistent arrangement. Based on the optimized grouping constraint, a connection-probability from pixels in nCRF to center pixel of CRF is calculated to decide the contour grouping significance of center pixel.

2.4.1. Connection-probability

According to rules of low curvature and local proximity, the connection-probability function should attenuate with the increasing of distance and curvature. So we define connection-probability between pixel A and B in Fig. 6 as follow.

$$P_{B \rightarrow A}(\theta_i, \theta_j) = \exp \left(-\frac{s^2 + \varepsilon_2 k^2}{2\sigma^2} \right) \quad (12)$$

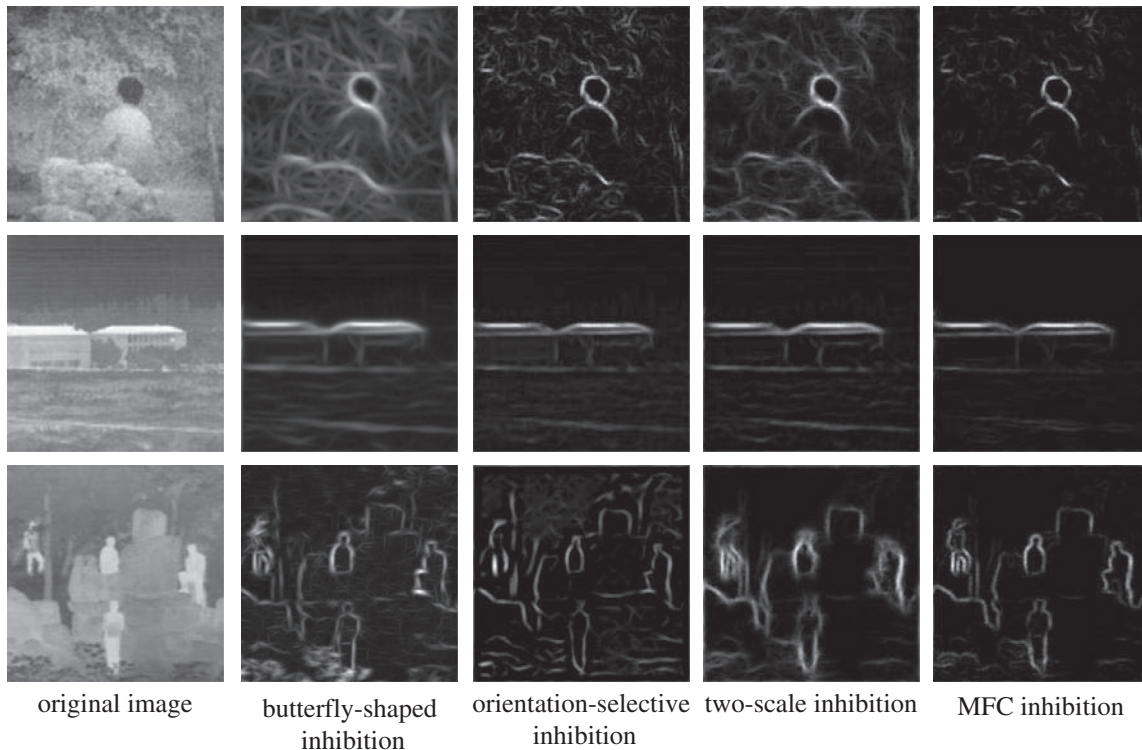


Fig. 5. Inhibition results by different methods. Parameters are set as $c_1 = 0.4$, $\sigma_\Delta = 0.5$, $\sigma = 3$ for LLL image (close shot LLL1 in first row) and $\sigma = 2.5$ for infrared image (distance infrared1 in second row and close shot infrared1 in third row).

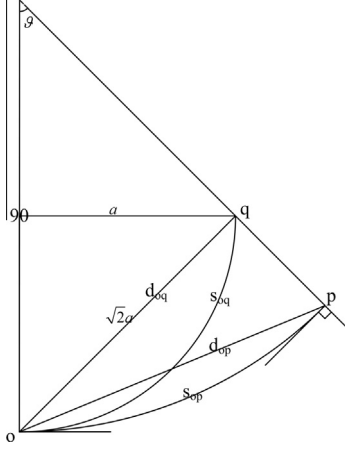


Fig. 7. Geometrical analysis of arc length and Euclidean distance connection.

Here $\vec{P}_{B \rightarrow A}(\theta_i, \theta_j)$ is the connection vector, $P_{B \rightarrow A}(\theta_i, \theta_j)$ is the value of connection-probability, connection orientation is θ_{AB} ($0 \leq \theta_{AB} < \pi$), $\varepsilon_2 > 0$ is the coefficient to adjust the proportion of distance and curvature in connection-probability, σ affects the size of connection field. To keep the criterion of curvature constancy, co-circularity rule is still considered. So the geometrical relationship between pixel A and B should fit Eq. (11). Then the

$$\begin{cases} \theta_j = 2\theta_{AB} - \theta_i \pm \delta, \varphi = (\theta_i - \theta_j)/2 & 0 \leq \theta_i < \pi/2, 0 \leq |\theta_i - \theta_{AB}| \leq \pi/4 \\ & \text{or } \pi/2 \leq \theta_i < \pi, 3\pi/4 \leq |\theta_i - \theta_{AB}| \leq \pi \\ \theta_j = 2\theta_{AB} - \theta_i + \pi \pm \delta, \varphi = \pi/2 + (\theta_i - \theta_j)/2 & 0 \leq |\theta_i - \theta_{AB}| \leq \pi/4 \\ \theta_j = 2\theta_{AB} - \theta_i - \pi \pm \delta, \varphi = \pi/2 - (\theta_i - \theta_j)/2 & 3\pi/4 \leq |\theta_i - \theta_{AB}| \leq \pi \end{cases} \quad (13)$$

spatial distance between pixel A and B is $d = \sqrt{m^2 + n^2} = 2r \sin \varphi$, arc length is $s = \varphi d / \sin \varphi$, and curvature is $k = 1/r = 2 \sin \varphi / d$.

Compared with [12], we use arc length instead of Euclidean distance to measure local proximity. Pixel in discontinuous contour is actually connected by curve along the tangent direction, so the connection distance between two points represented by arc length is more accurate than Euclidean distance.

As is shown in Fig. 7 ($\psi < \pi/4$), if we use the Euclidean distance to evaluate the local proximity, point O tends to connect with point Q for $d_{op} > d_{oq}$. If we use the arc length to evaluate the local proximity, O tends to connect with P for $s_{op} < s_{oq}$. According to the characteristic of low-curvature, O is more suitable to connect with P. Therefore the arc length connection is more accurate and accordant with Gestalt rules. Actually, in Fig. 7, there are $s = \psi r$, $d_{op} > d_{oq}$, $s_{oq} = \frac{\pi}{2} a$, $s_{op} = \psi(a + \frac{a}{\tan \psi}) = \psi a (\frac{\sin \psi + \cos \psi}{\sin \psi})$, so $s_{op} \geq s_{oq}$ when $\psi \geq \pi/4$ and $s_{op} < s_{oq}$ when $\psi < \pi/4$.

controls the proportion of curvature and arc length ($0 < \varepsilon_2 \leq 1$), the larger the ε_2 is, the flatter the excitatory connection is. The deviation angle $|\theta_i - \theta_{AB}|$ between $\vec{e}_{BA}$ and $\vec{e}_1$ (in Fig. 6) impacts the probability of excitatory connection, the larger the $|\theta_i - \theta_{AB}|$ is, the lower the excitatory connection probability is. By analyzing the relation from ε_2 and $|\theta_i - \theta_{AB}|$ to connection field, we can see that the connection is significant only in a range of deviation angle.

As a conclusion drawn in [12] that if the deviation angle $|\theta_i - \theta_{AB}|$ is more than $\pi/4$, it cannot guarantee that the curvature decreases with the orientation contrast. So it is constrained that $|\theta_i - \theta_{AB}| \leq \pi/4$. However, according to the numerical analysis, we convincingly explain that the constraint can be extended to $(|\theta_i - \theta_{AB}| \leq \pi/4) \cup (3\pi/4 \leq |\theta_i - \theta_{AB}| \leq \pi)$ in FC facilitation model.

In Fig. 8(a), when $\pi/2 < |\theta_i - \theta_{AB}| \leq \pi$ and $|\theta_i - \theta_{AB}| \leq \pi/2$, the connection field is full symmetry. When $|\theta_i - \theta_{AB}| \leq \pi/2$, the smaller the orientation contrast between center and surround is, the smaller the curvature of co-circularity points is. When $\pi/2 < |\theta_i - \theta_{AB}| \leq \pi$, the smaller the orientation contrast is, the larger the curvature is. In Fig. 8(b), when we restrain $\varepsilon_2 = 0.1$ and $P_{B \rightarrow A} \geq 0.3$ the maximum deviation angle is 43.6° . Regardless of the value of ε_2 there is barely reaction in the range of $\pi/4 < |\theta_i - \theta_{AB}| < 3\pi/4$, so the new constraint $(|\theta_i - \theta_{AB}| \leq \pi/4) \cup (3\pi/4 \leq |\theta_i - \theta_{AB}| \leq \pi)$ can both ensure the effective range of connection and reduce computing work. We assume that connection happens only when $0 \leq |\theta_i - \theta_{AB}| \leq \pi/4$ and $3\pi/4 \leq |\theta_i - \theta_{AB}| \leq \pi$, Eq. (11) is modified as follow.

Besides, compare the connection field of our facilitation model with model in [12] (Fig. 8(c)), connection in [12] tends to be more flat, whose connection field is similar to our model with $\varepsilon_2 = 1$. Then gaps may occur at corners with high-curvature. ε_2 is adjustable in FC facilitation model, so both smooth or high-curvature connection can be achieved. The value of ε_2 will be discussed below. As is shown in Fig. 9, it is better connected at inflexion points by our model.

Furthermore, if the surround deviates from the orientation axis of the center excessively, it will fail to form a co-aligned configuration. So we introduce a new constraint into the connection-probability, which is the projection from connection direction $\vec{e}_{BA}$ to preferred direction $\vec{e}_1$ at A. Then the connection-probability is amended as $P_{B \rightarrow A}(\theta_i, \theta_j) = \exp(-\frac{s^2 + \varepsilon_2 k^2}{2\sigma^2}) \cdot \cos(|\theta_i - \theta_{AB}|^+)$, $|\theta_i - \theta_{AB}|^+ = \min(|\theta_i - \theta_{AB}|, \pi - |\theta_i - \theta_{AB}|)$. The connection-probability is finally calculated as:

$$P_{B \rightarrow A}(\theta_i, \theta_j) = \begin{cases} \exp\left(-\frac{s^2 + \varepsilon_2 k^2}{2\sigma^2}\right) \cdot \cos(|\theta_i - \theta_{AB}|^+) & (0 < |\theta_i - \theta_{AB}| \leq \pi/4) \cup (3\pi/4 \leq |\theta_i - \theta_{AB}| < \pi) \\ \exp\left(-\frac{d^2}{2\sigma^2}\right) & (|\theta_i - \theta_{AB}| = 0) \cup (|\theta_i - \theta_{AB}| = \pi) \end{cases} \quad (14)$$

2.4.2. Constraint condition

Fig. 8 is the distribution field of the excitatory connection probability for short line (unit length), which is computed by Eq. (12). In Fig. 8, connection-probability attenuates with the length of smooth path increasing and tends to keep the direction of straight line. ε_2

2.4.3. Grouping significance

All points (e.g. pixel B in Fig. 6) in the connection field of pixel A will connect to A. The contour grouping significance of pixel A is

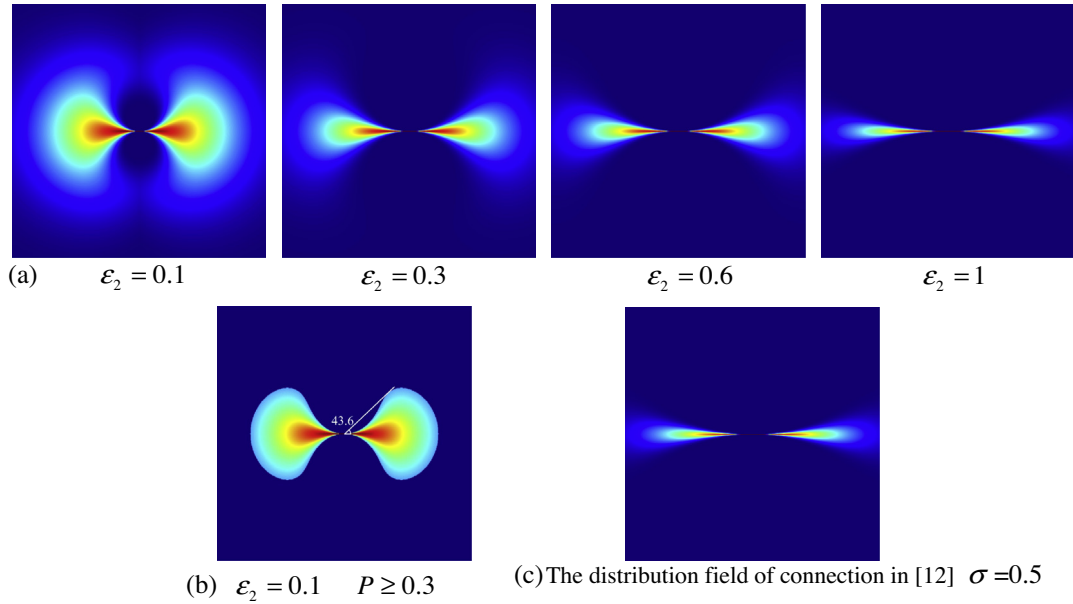


Fig. 8. The distribution field of the excitatory connection probability for short line. ($\sigma = 0.5$ $0 < \varepsilon_2 \leq 1$).

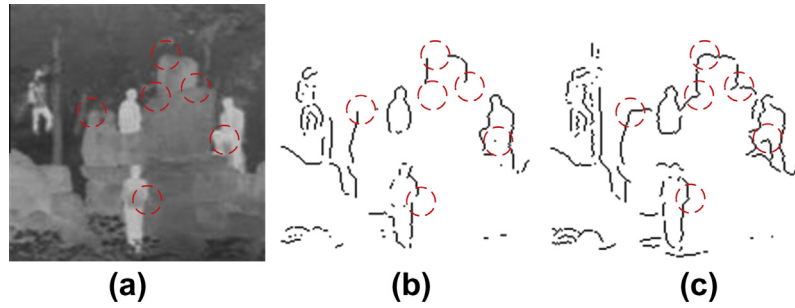


Fig. 9. Comparison of different facilitation model for contour extraction. (a) is an infrared image, (b) is facilitation model in [12], (c) is FC facilitation model in this paper ($\varepsilon_2 = 0.6$).

got by superposing all connection vectors at A. But only when the nCRF points meet the geometrical relationship in Eq. (13) with the center pixel in CRF, can they have probability to connect with center pixel.

In addition, the nCRF points should have prominent Gabor energy in their orientation θ_j to avoid mutual facilitation in homogeneous region. So the contour grouping significance $\omega_{fc}(x, y, \theta_i)$ of A is defined as the combination of optimized grouping constraint and local orientation energy of B.

$$\omega_{fc}(x, y, \theta_i) = \sum_{B \in S_{nc}, j \in [0, N_\theta)} P_{B \rightarrow A}(\theta_i, \theta_j) \cdot E_\sigma(B; \theta_j) \quad (15)$$

The existing facilitation models only consider the superposed connection from surrounding pixels to center pixel. That is only the situation for center pixel to be connected is calculated, without considering the attribute of center pixel itself. For region of background texture, the Gabor energy of each surrounding pixel in each direction is weak. For edge area of the target, the Gabor energy of surrounding pixel is strong only in part of directions. However, the facilitation effect is a superposed result, which is computed by accumulate the connection-probability from each surrounding pixel in each direction toward the center pixel. This could cause similar facilitation response in contour area and homogenous region.

In this paper we consider that if the central pixel is potential contour the facilitation response increases. On basis of the superposed connection of surrounding pixels, we consider the Gabor en-

ergy of center pixel. So the FC facilitation result depends on contour grouping significance and Gabor energy of center pixel. Accordingly, the function of FC facilitation $f_\sigma(x, y; \theta_i)$ can be calculated as follow.

$$f_\sigma(x, y; \theta_i) = \omega_{fc}(x, y, \theta_i) \cdot E_\sigma(x, y; \theta_i) \quad (16)$$

Center pixel itself with strong Gabor energy (potential contour) has a higher probability to be connected. Central pixel itself with weak Gabor energy (background texture) has a low probability to be connected. This can further distinguish between facilitation response of background and contour. As is shown in Fig. 10, compared with the blurry facilitation boundaries by [12], our model has more sharpened edges and the facilitation contour is more remarkable.

In night vision images, as for the pixels in internal texture region, they and their connection points are in random direction and have little Gabor energy because of the homogeneity, so they get weak facilitation response. The pixels in boundary connect mutually with large Gabor energy and consistent distribution to achieve high grouping significance and strong facilitation response. In Fig. 10, it is effective for FC facilitation to keep regular grouped contours and reduce noise interference. In addition, fuzzy connection will contribute to enhance the response of weak contours and fill in the gap between boundaries. Thus the FC facilitation model performs well in contour extraction and visuospatial integration.

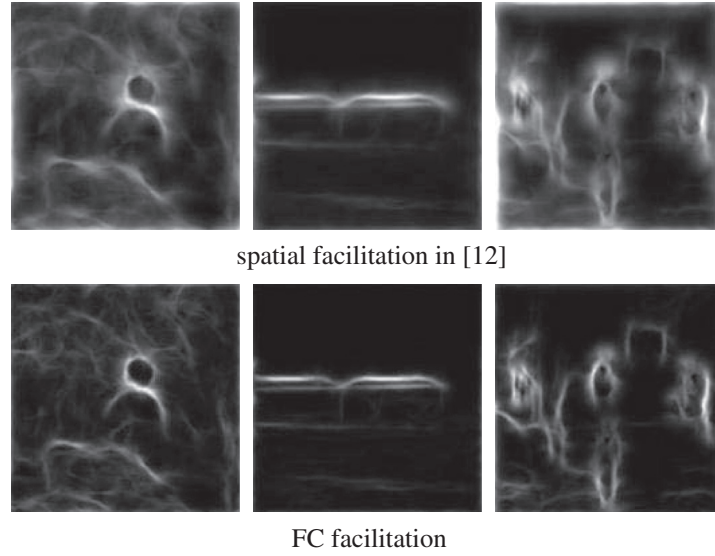


Fig. 10. Facilitation results in night vision images. Parameters are set as $e_2 = 0.6$, $\sigma = 3$ for LLL image and $\sigma = 2.5$ for infrared image.

2.5. Multi-scale iterative attention method

Physiology studies indicate that neuron in visual cortex accomplishes a particular visual task in a dynamic and flexible way [29]. When the peripheral region of CRF is long time stimulated, the CRF area will be multiple times larger [30,31]. Inspired by this biological characteristic, we propose a multi-scale iterative attention method to achieve dynamic modulation of nCRF.

The area of CRF depends on scale factor σ . Large-scale allows larger scope of interaction to extract contour of big target. But it may weaken contour of small target. Small-scale makes the modulation process more localized, which is conducive to keep small target. But it also retains some texture and noise. The proposed method increases the scale during iteration process to notice salient contour of targets in different sizes and inhibit background.

We define variable σ_t to replace σ in the proposed models, which increases linearly with the iteration number: $\sigma_t = \sigma_0 + \eta t$. In each step of iteration process, modulation result in each orientation is given as $R_{\sigma_t}(x, y, \theta_i)$. The final output is the maximum response over all orientations.

$$\frac{d}{dt} R_{\sigma_t}(x, y, \theta_i) = -R_{\sigma_t}(x, y, \theta_i) + |aR_{\sigma_t}(x, y, \theta_i) - bR_{\sigma_t}(x, y, \theta_i) + c f_{\sigma_t}(x, y, \theta_i)|^+ \quad (17)$$

$$R_{\sigma_t}(x, y) = \max \{R_{\sigma_t}(x, y; \theta_i) | i = 0, 1, \dots, N_\theta - 1\} \quad (18)$$

Here a , b and c control the proportion of inhibition and facilitation, the initial value of R_{σ_t} is set as E_{σ_0} .

In this multi-scale iterative attention method, the previous output at smaller scale is the input of next larger scale. It is difficult to control single scale adaptively to adapt for all targets and obtain the accurate and complete contour of target with any size. For small scale, it can inhibit homogeneous regions and retain contours of small target with fewer gaps. The facilitation is prone to obtain more natural and continuous gaps connection, but it remains a great deal of texture boundaries. For large scale, it has the ability to suppress texture to a greater degree and retain main contours, but it tends to lose small object contour and bring in big gaps, which is hard for facilitation to connect naturally and continuously.

Therefore, we operate the previous output at smaller scale with a larger scale and increase the scale in iteration. The target

contours are continuous and texture boundaries are scattered. So that small-scale inhibition makes the contour of small target be retained, small-scale facilitation makes the contour intact; large-scale keeps contour of large target and inhibit texture in a large extent. Meanwhile the continuous small target contours and scattered texture boundaries in surrounding are obvious different, which preserves contour of small target in large-scale inhibition. Combined with facilitation, the response of small target contour is growing through iteration process. Thus the iteration process has the ability to constantly suppress texture boundaries, enhance contour of small target also keep contour of large target. Increasing facilitation scale gradually will make discontinuous contours connection more natural and accurate.

As is shown in Fig. 11, in small-scale contour of small target is kept ((a) and (b) green¹ circle) with some residual texture, large-scale loses contour of small target ((c) and (d) green circle) but has less residual texture, which is an irreconcilable problem for contour extraction of targets in different sizes by single scale. However, small target is strengthened and large target is highlight in multi-scale iterative attention method ((e) green circle).

A Gaussian noise added circle is used to illustrate the effect of iteration in contour extraction, as is shown in Fig. 12. During the process of multi-scale analysis, MFC inhibition weakens homogeneous surrounding and irregular noises gradually to stand out main contour, FC facilitation strengthens excitatory connection to ensure the continuity of contour. The method performs well in extracting contour from serious noises.

3. Models analysis

3.1. For the facilitation model, why is the term “fuzzy connection” used? How was the “fuzzy” strategy employed during grouping contour segments?

First, the inhibition model may lead to contour disconnection. Contour discontinuity in serious condition must be connected. Second, if we analyze the grouping of discontinuous contour after suppression to determine the connection, the connection information is completely depends on the suppression outputs, leading to

¹ For interpretation of color in Figs. 11 and 13, the reader is referred to the web version of this article.

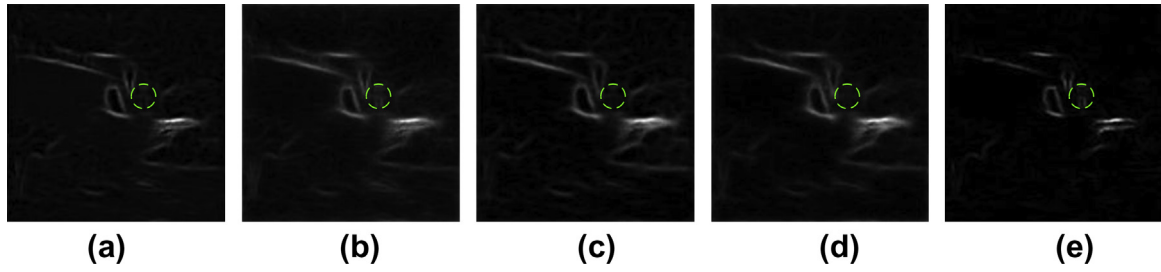


Fig. 11. Comparison of the small-scale model, large-scale model and the multi-scale iterative attention method. (a) and (b) are inhibition and inhibition + facilitation results in small scale. (c) and (d) are inhibition and inhibition + facilitation results in large scale. (e) is the iterative result.

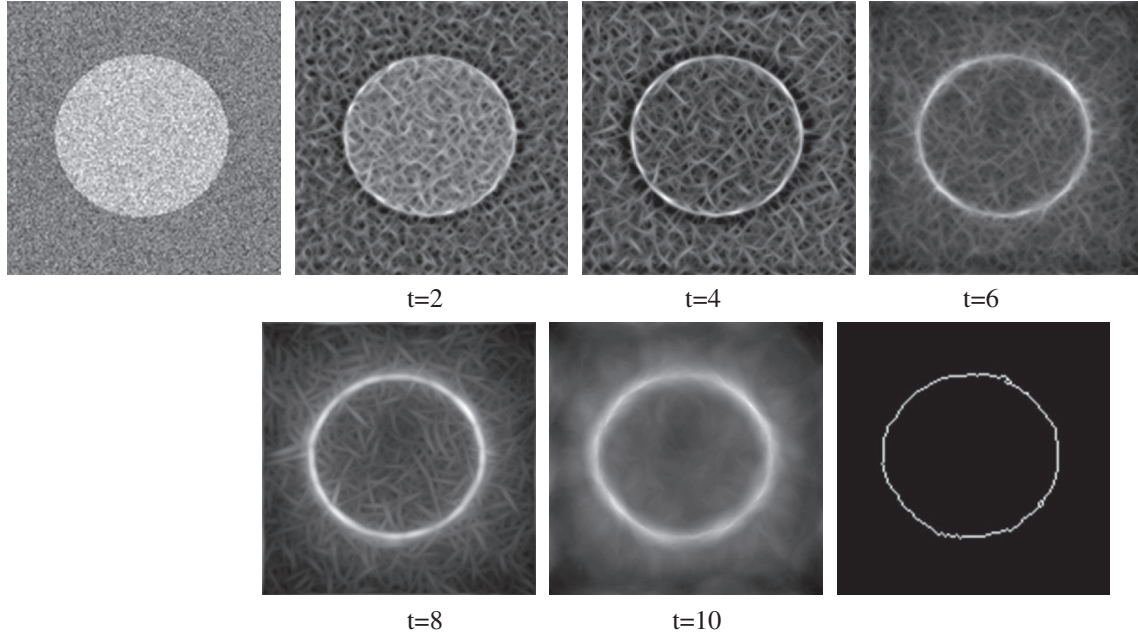


Fig. 12. Iterative attention results of the analog data. Parameters are set as $\sigma_0 = 0.5$, $\eta = 0.35$, $a = 1$, $b = 0.6$, $c = 0.4$, $t = 10$.

possible superposed errors. If we analyze the information of grouping facilitation directly on original image and calculate the probability of all potential contour pixels to be connected (It can be seen from the FC facilitation model that the potential contour segments have high probability to be connected), the facilitation effect is not dependent on the inhibition outputs. It will reduce superposed errors, which can improve the accuracy of contour extraction. Also inhibition and facilitation effects are independent, parallel processing can be achieved to improve the computational efficiency.

Fuzzy connection can realize the analysis of grouping facilitation on original image. From the formulas in FC facilitation model, we can see that any point in image can receive the “vote” from its neighborhood. Accumulation of the “vote” is the grouping connection probability at this point. So each point in image has not the only state to be connected or not, but each one is connecting or connected to a certain probability, which is the concept of “fuzzy connection”. The connection-probability is calculated for each point as “fuzzy” measure to assess its grouping significance, and is combined with inhibition results to enhance contour discrimination.

3.2. Experiment-based parameters analysis and discussion

In this paper, for MFC inhibition the proportionality coefficient and standard deviation of Gaussian distribution are set equal to $\varepsilon_1 = 0.4$ and $\sigma_\Delta = 0.5$. In FC facilitation the weighting of curvature

is set equal to $\varepsilon_2 = 0.6$. In multi-scale iterative attention the initial scale factor, growth rate of scale, iteration number and proportion of inhibition and facilitation are set as $\sigma_0 = 0.5$, $\eta = 0.35$, $t = 10$ and $a = 1$, $b = 0.6$, $c = 0.4$, respectively. It seems to involve many free parameters, but only three critical parameters affect the models' performance.

In contours output, non-maximum suppression and binaryzation is performed to remove false boundaries and select the salient contours with strong modulation response and long length. To discuss the influence of parameters, we use the definitions proposed by Grigorescu et al. [2]. c_t is the pixel set of truth salient contours; b_t is the pixel set of truth background edges; c_e is the pixel set of extracted contours; b_e is the pixel set of extracted background edges. Then correctly detected pixel set of contour is $c_e \cap c_t$; false negatives are given by $b_e \cap c_t$; false positives are given by $c_e \cap b_t$. We define three new measures in this paper: accurate rate $A_r = \text{num}(c_e \cap c_t) / \text{num}(c_t + b_t)$, false rate $F_r = \text{num}(b_e \cap c_t) / \text{num}(c_t + b_t) + \text{num}(c_e \cap b_t) / \text{num}(c_e \cap c_t)$, capacity of contour extraction $C = A_r - F_r$.

- (1) The weighting of curvature ε_2 decides the contour structure in FC facilitation model. The greater the curvature is, the smoother the contour is. Combined with Fig. 8 we discuss how the accurate rate A_r changes with ε_2 in Fig. 13(a). No matter what the value of ε_2 is, the straight line connection has the maximum probability to be executed. There are mainly straight line contours in distance infrared1 image

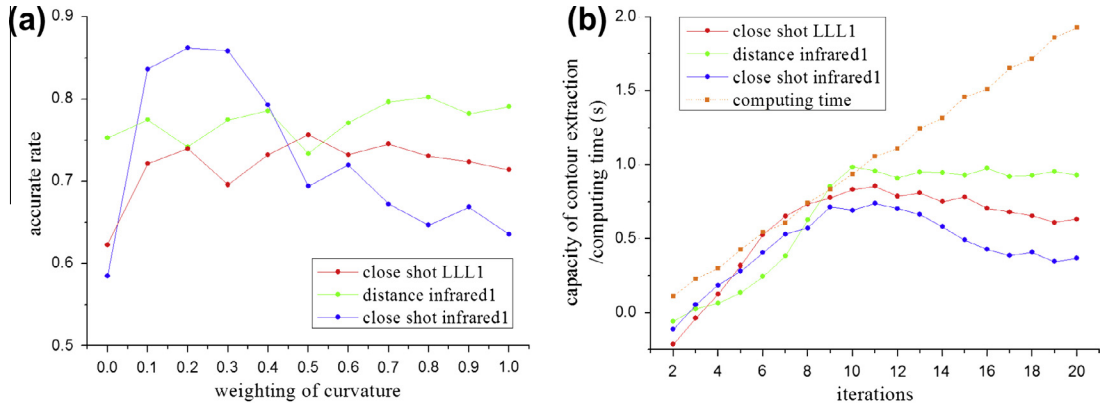


Fig. 13. Analyze the impact of parameters on models' performance. (a) The curve for accurate rate A_r changing with weighting of curvature ε_2 . (b) The curve for capacity of contour extraction C changing with iteration number t .

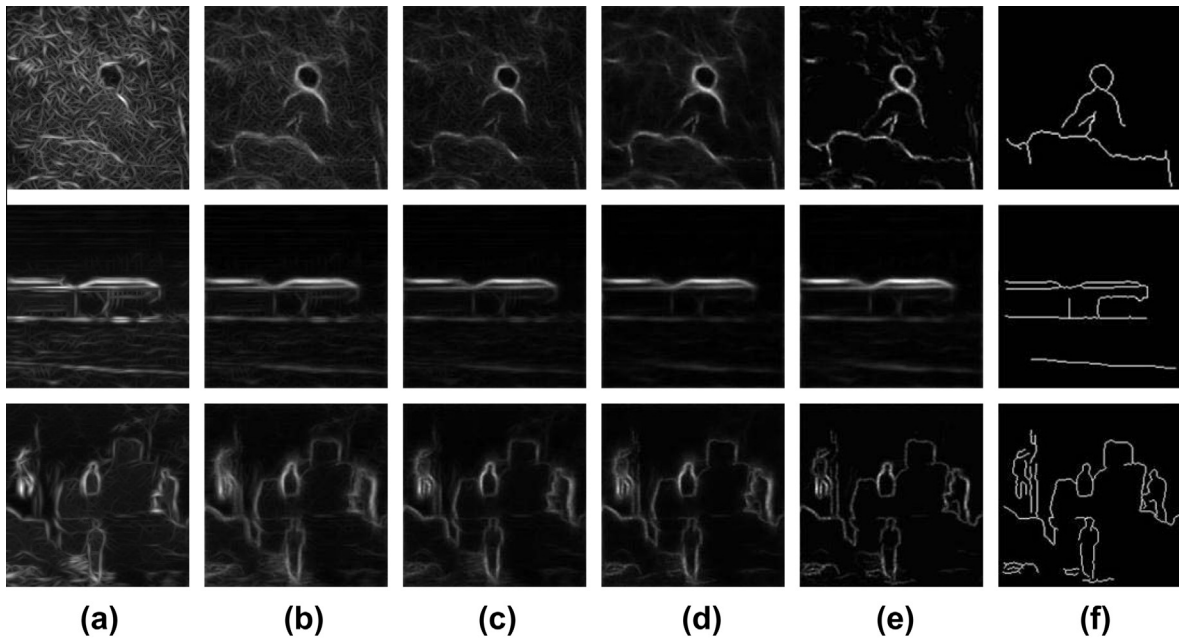


Fig. 14. The outputs of contour extraction in each step during iteration. (a) and (b) is the inhibition and inhibition + facilitation results in first iteration step. (c) and (d) is the inhibition and inhibition + facilitation results in sixth iteration step. (e) is the iterative result. (f) is the binaryzation output of contours. Parameters are set as $\sigma_0 = 0.5$, $\eta = 0.35$, $a = 1$, $b = 0.6$, $c = 0.4$, $t = 10$.

(image in second row of Fig. 5), so the larger the ε_2 is, the higher facilitation accuracy is (green line). There are many high-curvature connection points in close shot infrared1 image (image in third row of Fig. 5), making it achieves better facilitation result when ε_2 is small (blue line). There are mainly curve contour in close shot LLL1 (image in first row of Fig. 5), ε_2 ranges between the above two (red line). For the unknown night vision image, it is difficult to predict the contour structure, so we take a compromise in experiments and set the weighting as $\varepsilon_2 = 0.6$.

- (2) The iteration number t decides the accurate and false rate of contour extraction, also it determines the computing time of the multi-scale iterative attention method. We discuss how the capacity of contour extraction C changes with t ; in Fig. 13(b). When t is low, the modulatory coverage determined by scale factor is small and both contours and some textures are all retained, making A_r is high, but F_r is much greater, so $A_r - F_r$ is low. When t increases the scale is rising, response of part contours are weakened, meanwhile texture

inhibition is enhanced, leading to A_r and F_r both reduce. The combinatorial role of inhibition and facilitation makes the degree of texture inhibition large, but the degree of weakening in contour response is small. Thus the reducing of F_r is much rapider than A_r to gain a maximum value for $A_r - F_r$ in best iteration output. In subsequent iteration t further increases, the scale will be larger than many targets, so F_r is further reduced and A_r is also falling down. In this situation, residual textures are little so F_r reduces slowly. Despite the facilitation contours response are incessantly weakened due to mutual inhibition between each other, thus A_r decreases seriously. So $A_r - F_r$ is dropped. Besides, the larger t is, the greater the computing time is (orange line). As is shown in Fig. 13(b) (red, green and blue lines), when $t = 10$ the contour extraction obtains best performance in different image scenes.

- (3) The proportion of inhibition b and facilitation c also determine the performance of contour extraction. The response of targets may be weak when excessive inhibition in

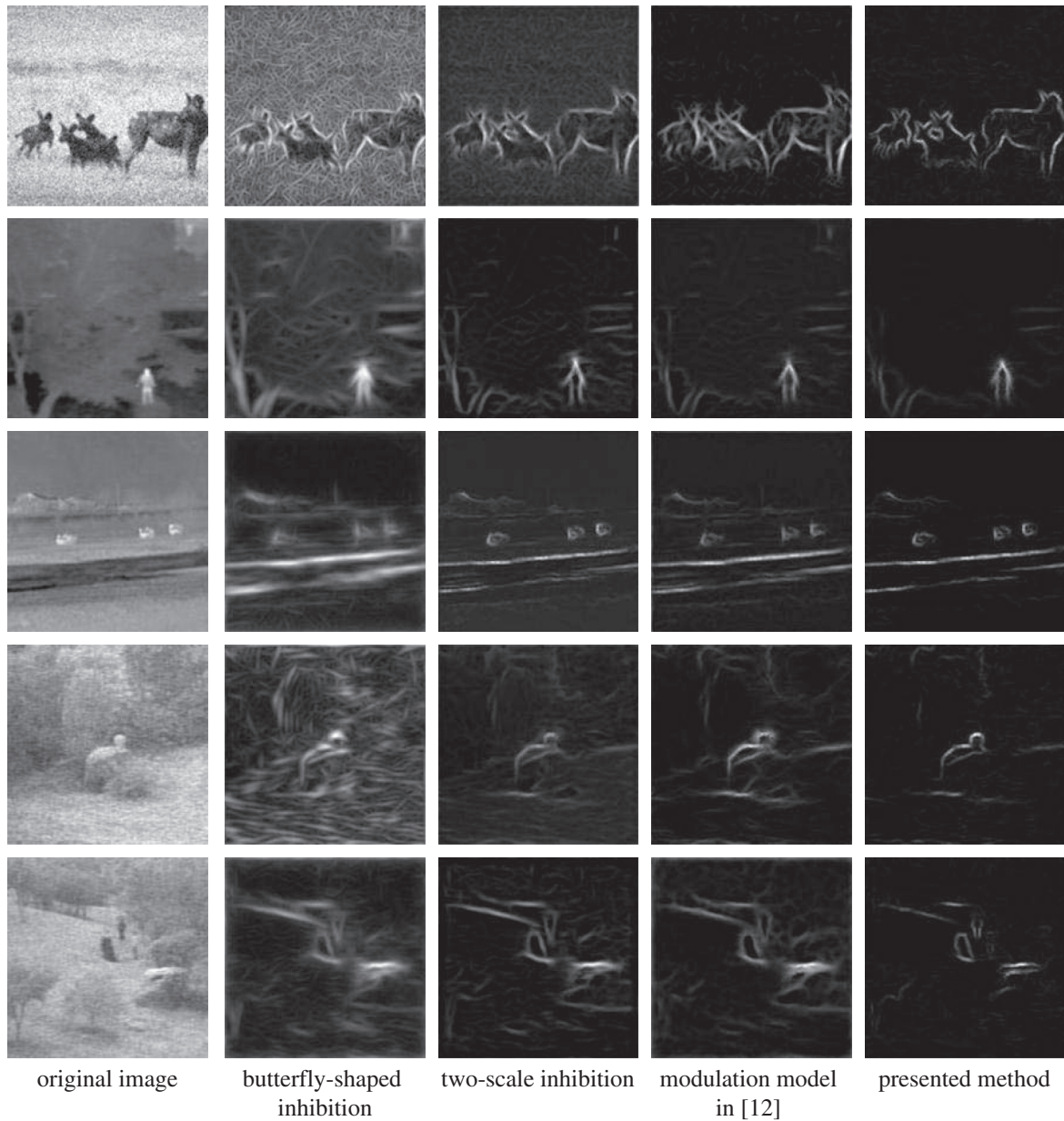


Fig. 15. Performance of contour extraction from night vision images in natural scenes by different methods. Rows from top to bottom are images of noisy hyena, close shot infrared2, distance infrared2, distance LLL1 and distance LLL2 respectively. Parameters are set as $\varepsilon_1 = 0.4$, $\sigma_\Delta = 0.5$, $\varepsilon_2 = 0.6$, $\sigma_0 = 0.5$, $\eta = 0.35$, $a = 1$, $b = 0.6$, $c = 0.4$, $t = 10$.

iteration process. Angular points may be ignored for the preference of smooth connection when facilitation plays a dominating role. So the value of b and c depends on the contour structure of significant objects in scene. Buildings scenes (many straight line contours) can set a larger proportion of inhibition, scenes with people and vehicles (complex shape of target contour) use a larger proportion of facilitation. We will further research on selective and adaptive modulation between inhibition and facilitation based on scene.

4. Empirical evaluation

Our method is applied to extract salient contour from night vision images. As is shown in Fig. 14, the output of small-scale inhibition has many boundaries of texture (Fig. 14(a)). The grouping significance of clutter texture is lower and grouping significance

of well-organized contour is much higher after inhibition + facilitation effect (Fig. 14(b)). Gradually increasing the iteration number to grow the scale, background in the output of small-scale is further weakened by inhibition in large-scale (Fig. 14(c)). The inhibition + facilitation role further enhances the response of salient contour (Fig. 14(d)). Finally Fig. 14(e) is the iteration result, Fig. 14(f) is the binaryzation output of contours. The results demonstrate the robustness of multi-scale analysis.

To further illustrate the property of the presented method, we compare the effect of our method with the butterfly-shaped inhibition, two-scale inhibition and modulation model in [12] on Gaussian noise added visible image, real LLL and infrared images in different natural scenes. As is shown in Fig. 15, despite of the different texture structure, noise distribution, scene complexity and target significance, the presented method can inhibit surrounding, extract contours of different targets with multi-size and maintain the integrity of contour better than the other models. So in nature

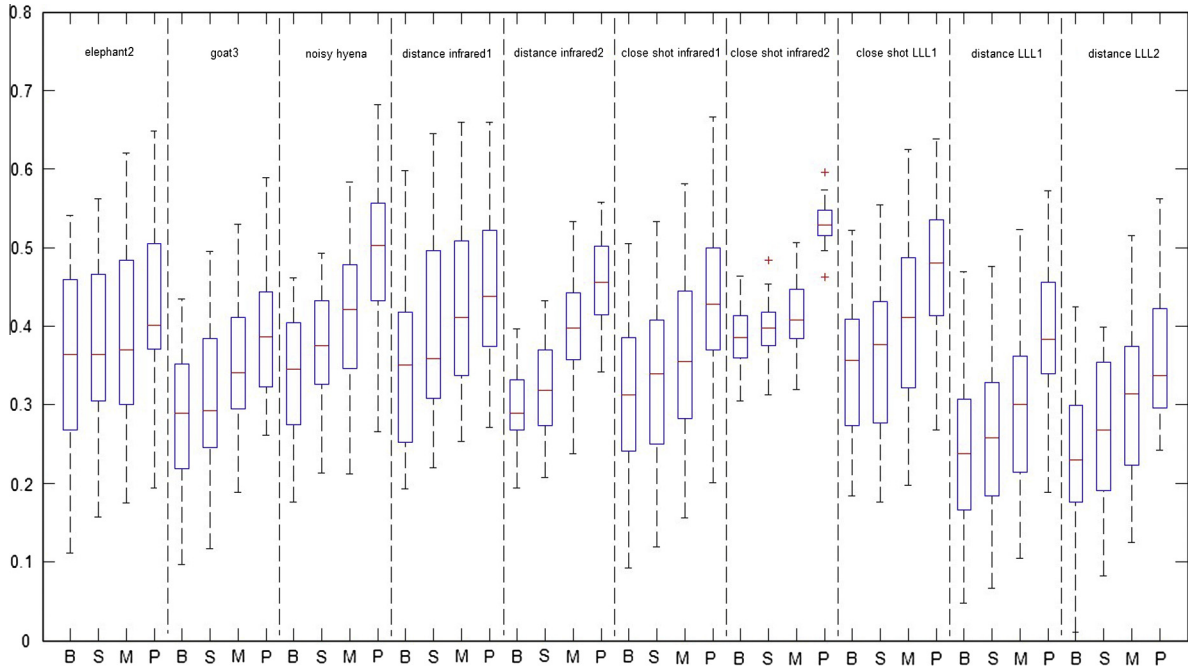


Fig. 16. Statistical box-and-whisker plots of the performance. B denotes the butterfly-shaped inhibition, S the two-scale inhibition, M the modulation model in [12], P the presented method.

scenes, our method performs well in separating edges of targets from heavy noises and abundant background and connecting discontinuous contour segments in LLL and infrared images to emphasize salient contours.

To test the robustness of the proposed model, we conducted a statistical analysis on the performance of 180 groups of different parameter combinations mentioned above ($\varepsilon_2 = [0:0.2:1]$, $t = [2:2:20]$, $b/c = [0.5, 1, 2]$) to calculate box-and-whisker plots for the butterfly-shaped inhibition, two-scale inhibition, modulation model in [12] and presented model.

The performance of contour extraction defined in [2] is $P = \text{num}(c_e \cap c_t) / (\text{num}(c_e \cap c_t) + \text{num}(b_e \cap c_t) + \text{num}(c_e \cap b_t))$. It is used to quantitatively evaluate the contour extraction. Fig. 16 shows this statistical box-and-whisker plot for ten images mentioned above in this study. In each plot, the top end of a whisker represents the best performance among the 180 different results for each image, and the horizontal red line in the box of a whisker shows the median value of 180 performance measures. The top and bottom edges of each box reflect the upper and lower median, respectively.

For all test images, the top end of whisker of the presented model is obviously higher than that of other models. More importantly, the presented model produces a significantly higher median performance measure, and the box height (i.e., the range between the upper and lower medians) of the presented model is smaller than that of most boxes of other models, which is particularly obvious in real LLL and infrared images. It reveals that our model performs better and has higher robustness in extracting salient contour from cluttered scene in night vision images.

5. Conclusions

In this work, according to characteristics of night vision images, we improve the receptive field models based on perception mechanism and propose an effective contour extraction method to suppress noise and texture and detect salient contour from LLL and infrared images in natural scene.

The optimized modulation method based on compound receptive field coordinates two aspects of surround inhibition and grouping facilitation. As for surround inhibition, we first bring forward a multi-feature contrast weighted inhibition model. Different from the orientation-selective mechanism, we focus on features in LLL and infrared images to evaluate the correlation of center and surround in feature space and weighting inhibition by the multi-feature difference. The MFC weighted inhibition model can eliminate details effectively and lower the possibility of mutual inhibition among contour elements. As for grouping facilitation, we innovatively present a fuzzy connection facilitation Model. Other from co-circularity method, we establish grouping facilitation with more comprehensive perception constraints to calculate connection-probability between center and surround and estimate grouping significance for each pixel. The FC facilitation model is consistent with Gestalt rule to achieve precise contour connection. Finally, we design a visual inspired multi-scale iterative attention method to accomplish hierarchical attention from details to main contours by changing scale factor in iteration process dynamically.

However, the response of targets may be weak due to the inhibition mechanism in iteration process. Angular points may be ignored for the preference of smooth connection when facilitation plays a dominating role. So we will research on selective and adaptive modulation between inhibition and facilitation in future work, that is how to decide the value of a , b and c intelligently.

Acknowledgements

This work was supported by the Natural Science Foundations of China (61231014 and 61373061).

Appendix A. Computation of multiple features

Consider a given local image of size $M \times N$, where M is number of rows and N is number of columns. The spatial frequency SF is defined in [32]. The row and column frequencies are given by $\text{Row.Freq} = \sqrt{\frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} [F(i,j) - F(i,j-1)]^2}$ and $\text{Column.Freq} =$

$\sqrt{\frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} [F(i,j) - F(i-1,j)]^2}$. The total frequency is then $SF = \sqrt{(\text{Row.Freq})^2 + (\text{Column.Freq})^2}$. This definition of frequency in the spatial domain indicates the overall activity level in an image.

Contrast *CON* measures the local variations in center and surround windows, which is actually defined as the local standard deviation $CON = \sqrt{\frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} [F(i,j) - \mu]^2}$, where μ is the local gray average.

Co-occurrence statistics evaluate the joint probability occurrence of the specified pixel pairs in gray-level co-occurrence matrix GLCM. In this paper, we consider GLCM in four directions (0,45,90,135) with offset of two pixels.

Co-occurrence statistics of correlation *cooCOR* measures the similarity of elements in rows or columns in GLCM, which can reflect the local gray correlation in image. *cooCOR* is calculated as $cooCOR = \sum_{i=1}^k \sum_{j=1}^k \frac{(ij)G(i,j) - u_i u_j}{s_i s_j}$, where G is the GLCM matrix, k is row and column number of G , $u_i = \sum_{j=1}^k \sum_{j=1}^k i \cdot G(i,j)$, $u_j = \sum_{i=1}^k \sum_{j=1}^k j \cdot G(i,j)$, $s_i^2 = \sum_{j=1}^k \sum_{j=1}^k G(i,j)(i - u_i)^2$, $s_j^2 = \sum_{i=1}^k \sum_{j=1}^k G(i,j)(j - u_j)^2$. *cooCOR* is big when the elements are uniform in GLCM and is small when the elements vary widely in GLCM. So it evaluates the consistency of image texture.

Co-occurrence statistics of homogeneity *cooH* measures the closeness of the distribution of elements in GLCM to the diagonal. *cooH* is calculated as $cooH = \sum_{i=1}^k \sum_{j=1}^k \frac{G(i,j)}{1 + (i-j)^2}$. Big diagonal elements in GLCM lead to big *cooH*, then image region with consistent gray has large *cooH*. So *cooH* evaluates the local homogeneity of image texture.

References

- [1] H.E. Jones, K.L. Grieve, W. Wang, A.M. Sillito, Surround suppression in primate V1, *J. Neurophysiol.* 86 (2001) 2011–2028.
- [2] C. Grigorescu, N. Petkov, M.A. Westenberg, Improved contour detection by non-classical receptive field inhibition, *Biol. Motivated Comput. Vision Second Int. Workshop* (2002) 50–59.
- [3] C. Grigorescu, N. Petkov, M.A. Westenberg, Contour detection based on nonclassical receptive field inhibition, *IEEE Trans. Image Process.* 12 (7) (2003) 729–739.
- [4] C. Grigorescu, N. Petkov, M.A. Westenberg, Contour and boundary detection improved by surround suppression of texture edges, *Image Vis. Comput.* 22 (2004) 609–622.
- [5] G. Papari, P. Campisi, N. Petkov, A. Neri, A biologically motivated multiresolution approach to contour detection, *EURASIP J. Adv. Signal Process.* 71828 (2007) 1–28.
- [6] N. Sang, Q.L. Tang, T.X. Zhang, Contour detection based on inhibition of primary visual cortex, *J. Infrared Millin. Waves* 26 (1) (2007) 47–60.
- [7] M. Ursino, G.E. La Cara, A model of contextual interactions and contour detection in primary visual cortex, *Neural Networks* 17 (2004) 719–735.
- [8] G.E. La Cara, M. Ursino, A model of contour extraction including multiple scales, flexible inhibition and attention, *Neural Networks* 21 (2008) 759–773.
- [9] C. Zeng, Y.J. Li, C.Y. Li, Center-surround interaction with adaptive inhibition: A computational model for contour detection, *NeuroImage* 55 (2011) 49–66.
- [10] C. Zeng, Y.J. Li, K.F. Yang, C.Y. Li, Contour detection based on a non-classical receptive field model with butterfly-shaped inhibition subregions, *Neurocomputing* 74 (2011) 1527–1534.
- [11] C.Y. Li, W. Li, Extensive integration field beyond the classical receptive field of cat's striate cortical neurons-classification and tuning properties, *Vision Res.* 34 (18) (1994) 2337–2355.
- [12] Q.L. Tang, N. Sang, T.X. Zhang, Extraction of salient contours from cluttered scenes, *Pattern Recogn.* 40 (2007) 3100–3109.
- [13] Q.L. Tang, N. Sang, T.X. Zhang, Extraction of salient contours via excitatory-inhibitory interactions in the visual cortex, in: 7th Asian Conference on Computer Vision (2006) 683–691.
- [14] L.R. Williams, K.K. Thornber, A comparison of measures for detecting natural shapes in cluttered backgrounds, *Int. J. Comput. Vision* 34 (2–3) (1999) 81–96.
- [15] W.S. Geisler, J.S. Perry, B.J. Super, D.P. Gallogly, Edge co-occurrence in natural images predicts contours grouping performance, *Vision Res.* 41 (6) (2001) 711–724.
- [16] A. Krupnik, L.A. Johnston, Contour grouping with prior models, *IEEE Trans. Pattern Anal. Mach. Intell.* 25 (6) (2003) 661–674.
- [17] S. Wang, T. Kubota, J.M. Siskind, Salient boundary detection using ratio contour, *Adv. Neural Information Process. Syst.* 16 (3) (2003) 231–240.
- [18] S. Wang, T. Kubota, J.M. Siskind, J. Wang, Salient closed boundary extraction with ratio contour, *IEEE Trans. Pattern Anal. Mach. Intell.* 27 (4) (2005) 546–561.
- [19] T.D. Nguyen, G. Lee, Color image segmentation using tensor voting based color clustering, *Pattern Recogn. Lett.* 33 (2012) 605–614.
- [20] J. Lim, J. Park, G.G. Medioni, Text segmentation in color images using tensor voting, *Image Vis. Comput.* 25 (2007) 671–685.
- [21] T.D. Nguyen, J. Park, G. Lee, Tensor voting based text localization in natural scene images, *IEEE Signal Process. Lett.* 17 (7) (2010) 639–642.
- [22] W. Chan, G. Coghill, Text analysis using local energy, *Pattern Recogn.* 34 (2001) 2523–2532.
- [23] F. Sengpiel, A. Sen, C. Blakemore, Characteristics of surround inhibition in cat area 17, *Exp Brain Res.* 116 (2) (1997) 216–228.
- [24] J.B. Levitt, J.S. Lund, Contrast dependence of contextual effects in primate visual cortex, *Nature* 387 (6628) (1997) 73–76.
- [25] G.A. Walker, I. Ohzawa, R.D. Freeman, Asymmetric suppression outside the classical receptive field of the visual cortex, *J. Neurosci.* 19 (23) (1999) 10536–10553.
- [26] U. Polat, K. Mizobe, M.W. Pettet, T. Kasamatsu, A.M. Norcia, Collinear stimuli regulate visual responses depending on cell's contrast threshold, *Nature* 391 (1998) 580–584.
- [27] M.K. Kapadia, G. Westheimer, C.D. Gilbert, Spatial distribution of contextual interactions in primary visual cortex and in visual perception, *J. Neurophys.* 84 (4) (2000) 2048–2062.
- [28] J. Xing, D.J. Heeger, Measurement and modeling of center-surround suppression and enhancement, *Vision Res.* 41 (5) (2001) 571–583.
- [29] V. Dragoi, M. Sur, Dynamic properties of recurrent inhibition in primary visual cortex: contrast and orientation dependence of contextual effects, *J. Neurophys.* 83 (2) (2000) 1019–1030.
- [30] C.D. Gilbert, T.N. Wiesel, The influence of contextual stimuli on the orientation selectivity of cells in primary visual cortex of the cat, *Vision Res.* 30 (11) (1990) 1689–1701.
- [31] M.W. Pettet, C.D. Gilbert, Dynamic changes in receptive-field size in cat primary visual cortex, *Proc. Natl. Acad. Sci. USA* 89 (1992) 8366–8370.
- [32] A.M. Eskicioglu, P.S. Fisher, Image quantity measures and their performance, *IEEE Trans. Commun.* 43 (12) (1995) 2959–2965.